7th International Protégé Conference
2nd Workshop on Visualizing Information
in Knowledge Engineering (VIKE'04)

Bethesda, MD July 6, 2004

Biomedical Knowledge Visualization

Olivier Bodenreider

Lister Hill National Center
for Biomedical Communications
Bethesda, Maryland - USA

UMLS Semantic Navigator

SemNav

http://umlsks.nlm.nih.gov*

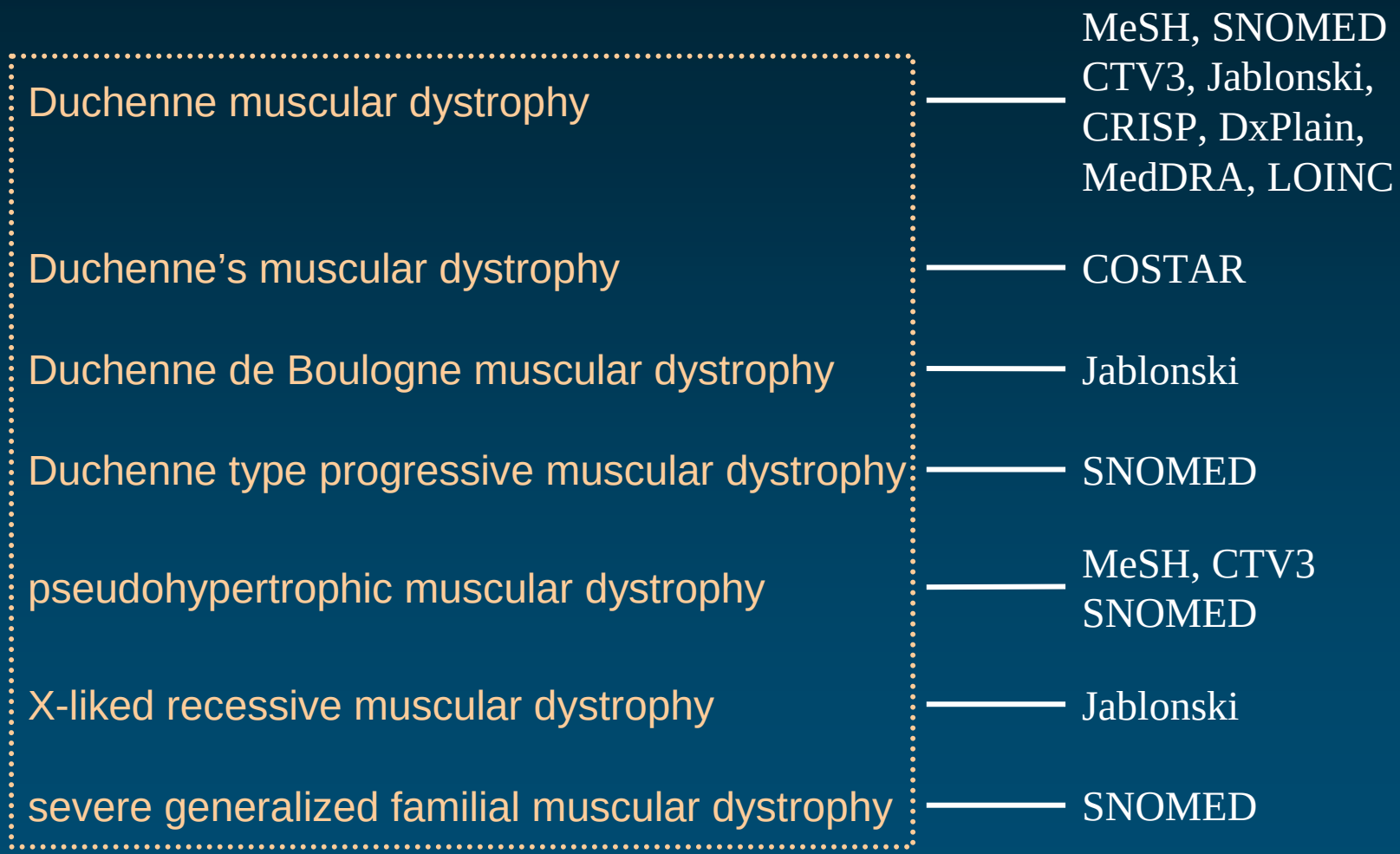
► SN Resources ► Semantic Navigator

(* free UMLS registration required)

Unified Medical Language System®

- Developed at NLM since 1990
- 15th edition in 2004
- Integrates some 60 terminological resources
  - Clinical vocabularies (including specialties)
  - Core terminologies (anatomy, drugs, med. devices)
  - Administrative terminologies, standards
- Integration
  - Synonymous terms are clustered in a concept
  - Hierarchies (trees) are combined in a graph structure

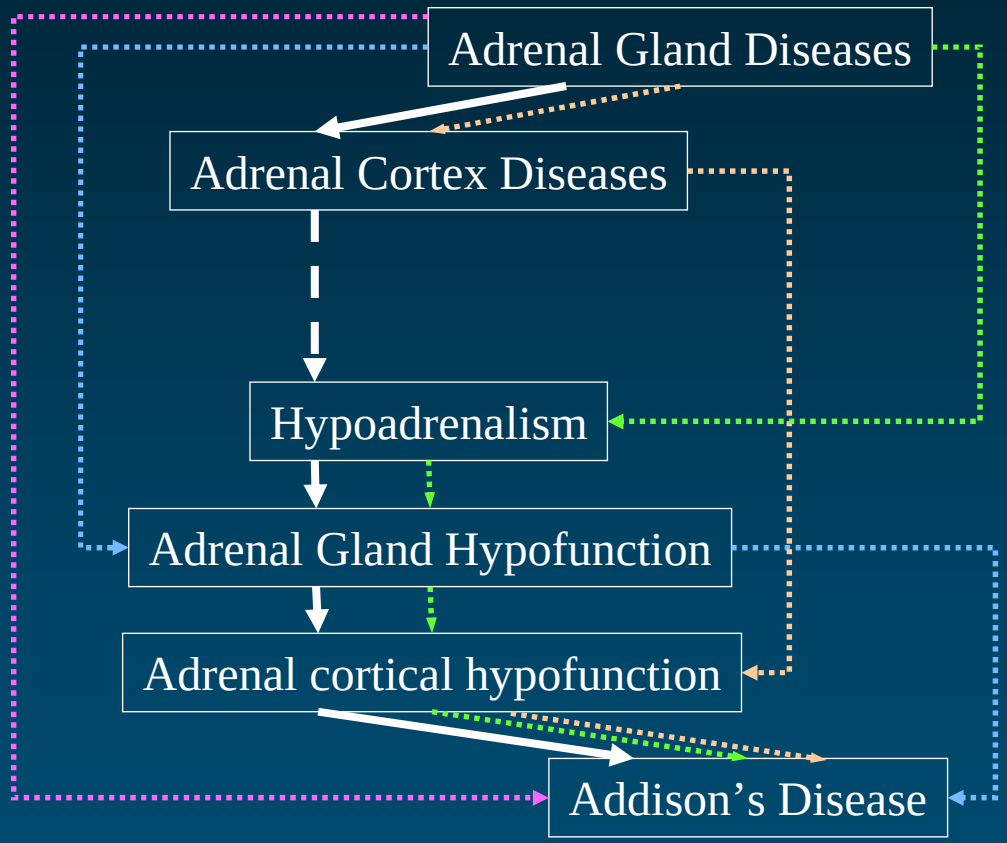
Terminology integration Terms



Terminology integration
Relationships

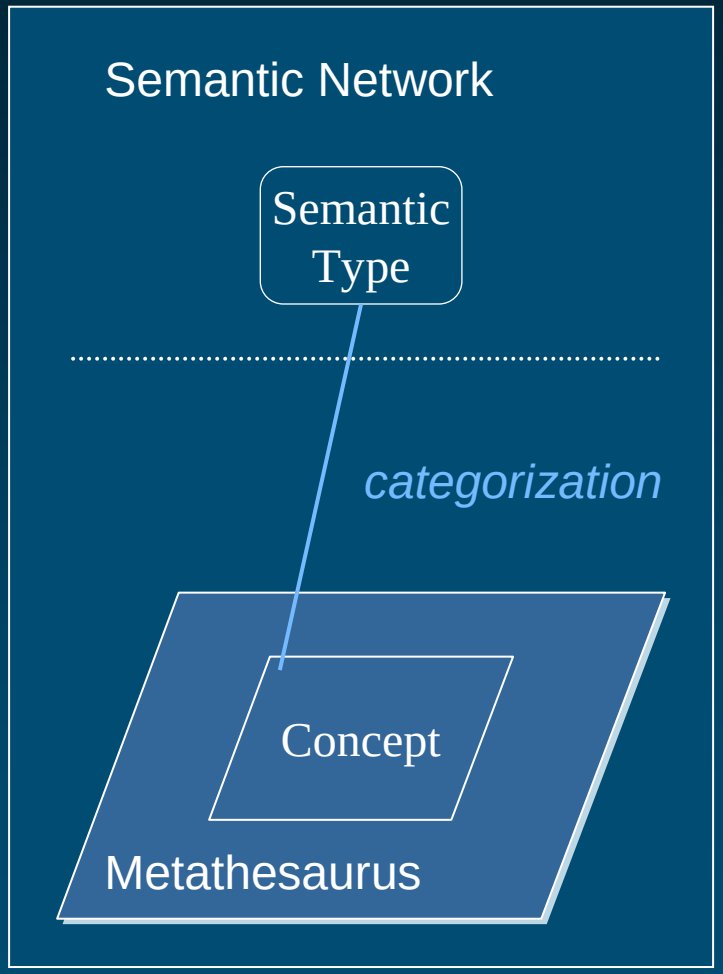
SNOMED
MeSH
AOD
Read Codes

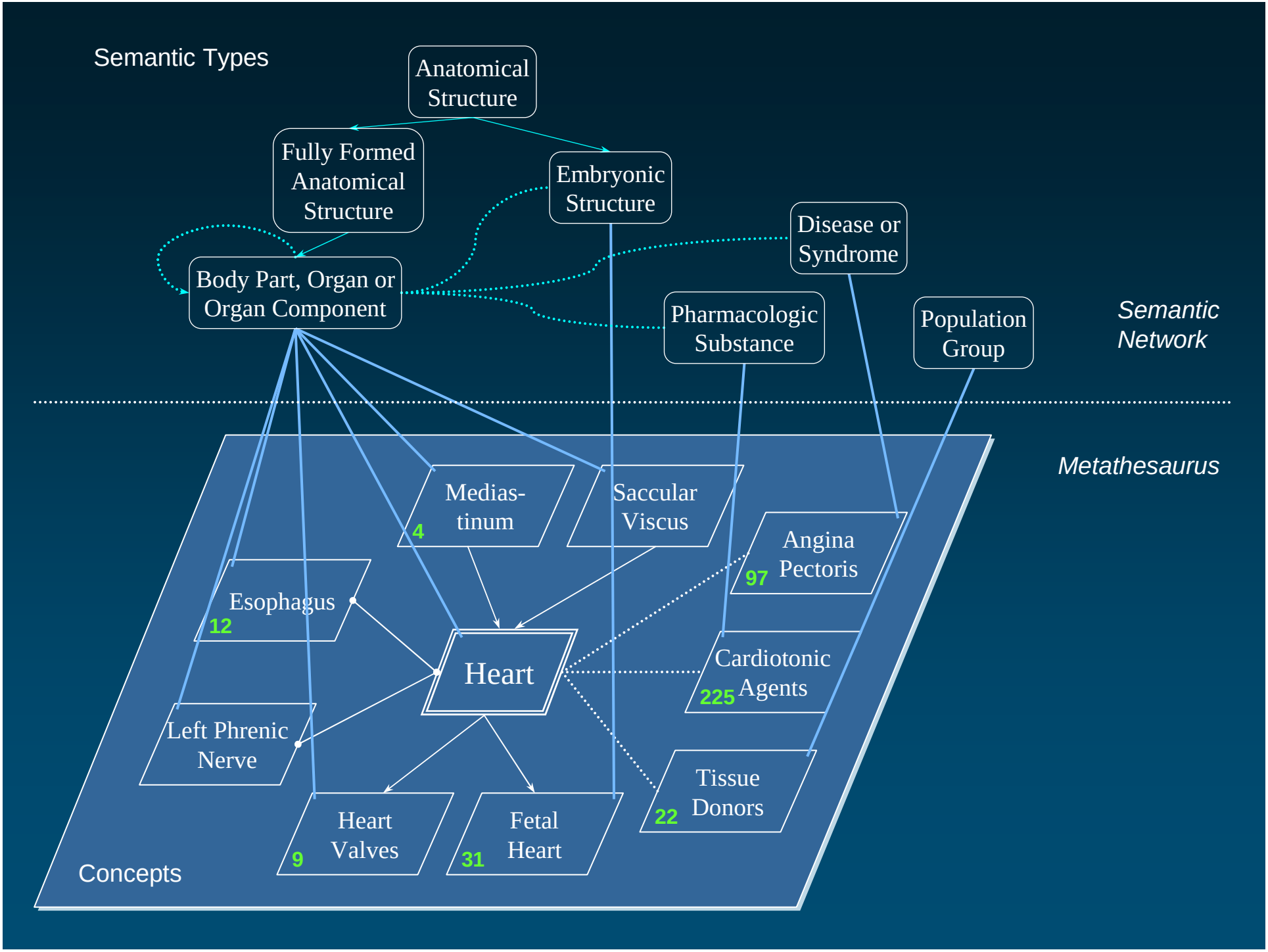
UMLS

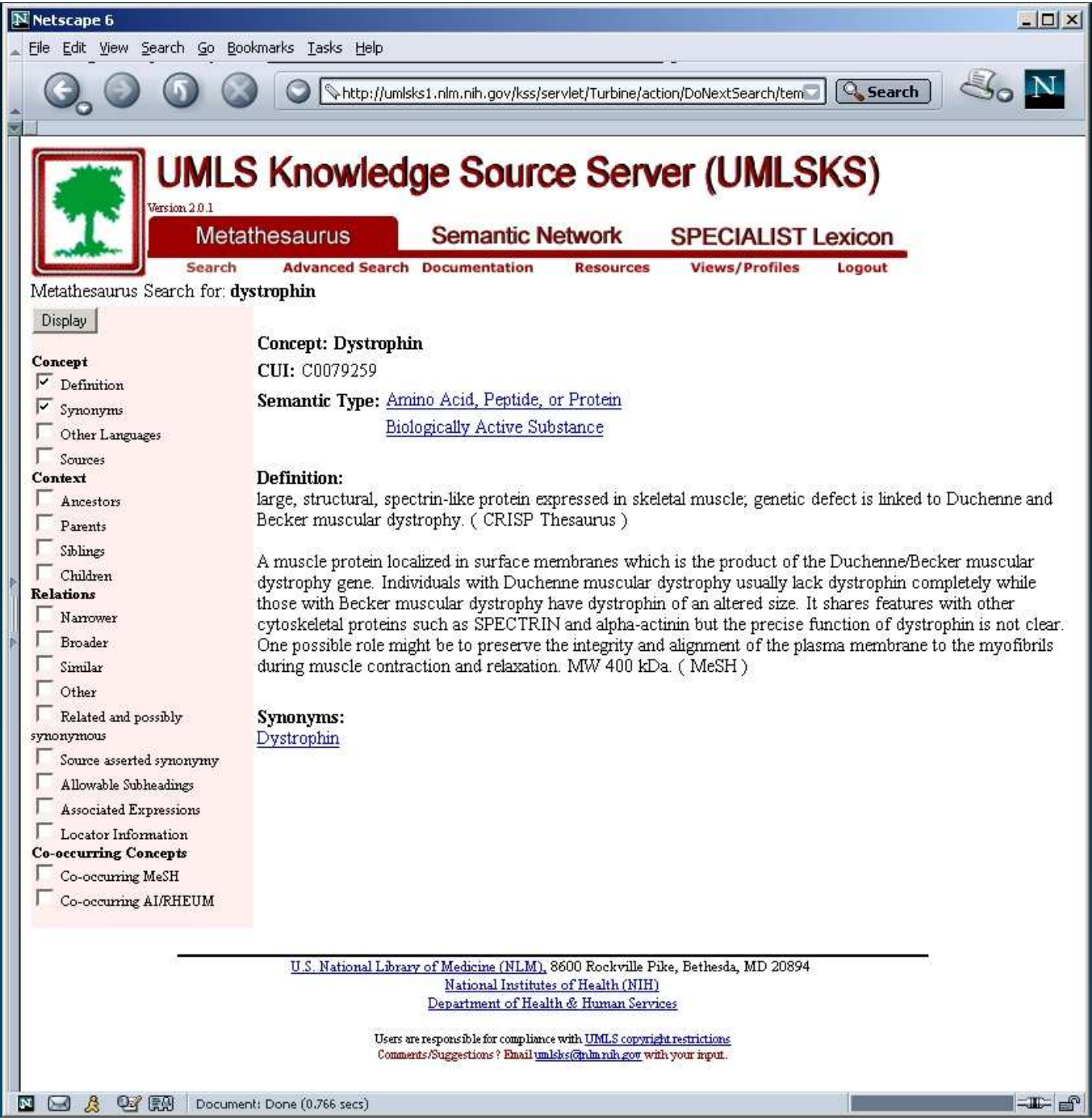


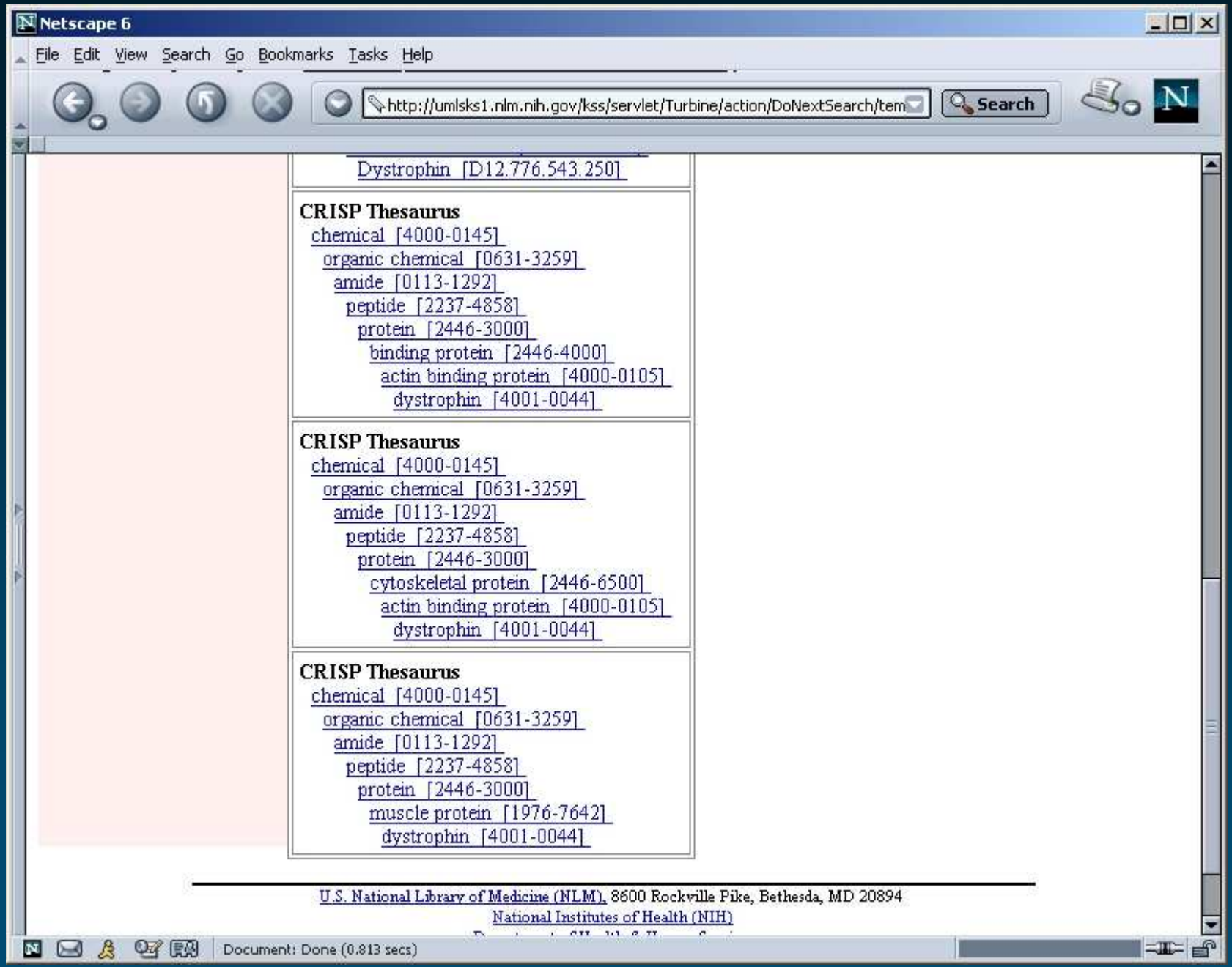
UMLS

- Two-level structure
  - Semantic Network
    - 135 Semantic Types (STs)
    - 54 types of relationships among STs
  - Metathesaurus
    - >1M concepts
    - ~12 M inter-concept relationships
  - Link = categorization




Semantic Types
Anatomical Structure
Fully Formed Anatomical Structure
Embryonic Structure
Disease or Syndrome
Body Part, Organ or Organ Component
Pharmacologic Substance
Population Group
Semantic Network
Metathesaurus
Medias-tinum
4
Saccular Viscus
Angina Pectoris
97
Esophagus
12
Heart
Cardiotonic Agents
225
Left Phrenic Nerve
Tissue Donors
22
Heart Valves
9
Fetal Heart
31
Concepts


Netscape 6
File Edit View Search Go Bookmarks Tasks Help
http://umlsks1.nlm.nih.gov/kss/servlet/Turbine/action/DoNextSearch/tem
Search
UMLS Knowledge Source Server (UMLSKS)
Version 2.0.1
Metathesaurus
Semantic Network
SPECIALIST Lexicon
Search
Advanced Search
Documentation
Resources
Views/Profiles
Logout
Metathesaurus Search for: dystrophin
Display
Concept
Definition
Synonyms
Other Languages
Sources
Context
Ancestors
Parents
Siblings
Children
Relations
Narrower
Broader
Similar
Other
Related and possibly synonymous
Source asserted synonymy
Allowable Subheadings
Associated Expressions
Locator Information
Co-occurring Concepts
Co-occurring MeSH
Co-occurring AI/RHEUM
Concept: Dystrophin
CUI: C0079259
Semantic Type: Amino Acid, Peptide, or Protein
Biologically Active Substance
Definition:
large, structural, spectrin-like protein expressed in skeletal muscle; genetic defect is linked to Duchenne and Becker muscular dystrophy. (CRISP Thesaurus)
A muscle protein localized in surface membranes which is the product of the Duchenne/Becker muscular dystrophy gene. Individuals with Duchenne muscular dystrophy usually lack dystrophin completely while those with Becker muscular dystrophy have dystrophin of an altered size. It shares features with other cytoskeletal proteins such as SPECTRIN and alpha-actinin but the precise function of dystrophin is not clear. One possible role might be to preserve the integrity and alignment of the plasma membrane to the myofibrils during muscle contraction and relaxation. MW 400 kDa. (MeSH)
Synonyms:
Dystrophin
U.S. National Library of Medicine (NLM), 8600 Rockville Pike, Bethesda, MD 20894
National Institutes of Health (NIH)
Department of Health & Human Services
Users are responsible for compliance with UMLS copyright restrictions
Comments/Suggestions ? Email umlsks@nlm.nih.gov with your input.
Document: Done (0.766 secs)


Netscape 6
File Edit View Search Go Bookmarks Tasks Help
http://umlsks1.nlm.nih.gov/kss/servlet/Turbine/action/DoNextSearch/tem
Search
Dystrophin [D12.776.543.250]
CRISP Thesaurus
chemical [4000-0145]
organic chemical [0631-3259]
amide [0113-1292]
peptide [2237-4858]
protein [2446-3000]
binding protein [2446-4000]
actin binding protein [4000-0105]
dystrophin [4001-0044]
CRISP Thesaurus
chemical [4000-0145]
organic chemical [0631-3259]
amide [0113-1292]
peptide [2237-4858]
protein [2446-3000]
cytoskeletal protein [2446-6500]
actin binding protein [4000-0105]
dystrophin [4001-0044]
CRISP Thesaurus
chemical [4000-0145]
organic chemical [0631-3259]
amide [0113-1292]
peptide [2237-4858]
protein [2446-3000]
muscle protein [1976-7642]
dystrophin [4001-0044]
U.S. National Library of Medicine (NLM), 8600 Rockville Pike, Bethesda, MD 20894
National Institutes of Health (NIH)
Document: Done (0.813 secs)

MeSH Browser

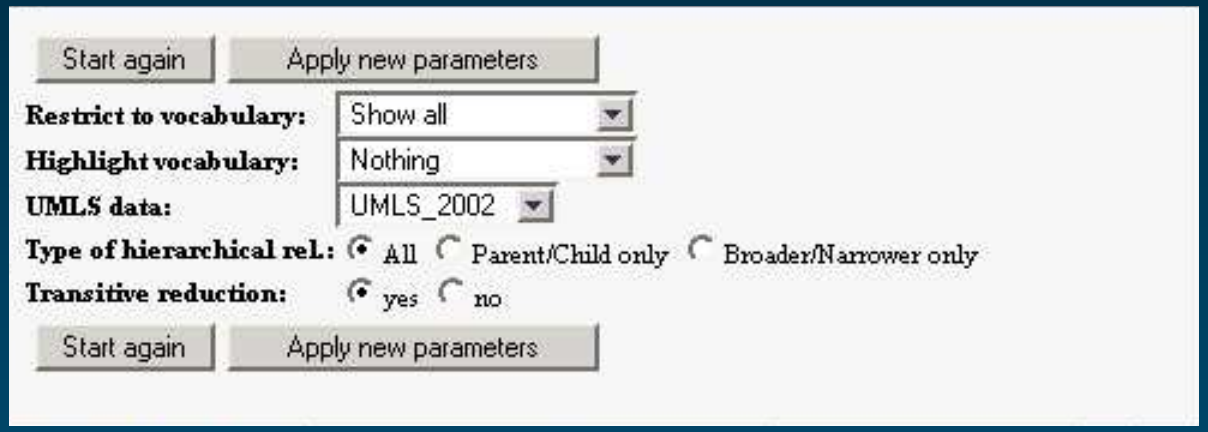
Amino Acids, Peptides, and Proteins [D12]
- Proteins [D12.776]
  - Contractile Proteins [D12.776.210]
    - Muscle Proteins [D12.776.210.500]
      - Actinin [D12.776.210.500.095]
      - Actins [D12.776.210.500.100]
      - Actomyosin [D12.776.210.500.154]
      - Calsequestrin [D12.776.210.500.220]
      - Dystrophin [D12.776.210.500.250]
      - Myogenic Regulatory Factors [D12.776.210.500.570] +
      - Myoglobin [D12.776.210.500.588]
      - Myosins [D12.776.210.500.600] +
      - Parvalbumins [D12.776.210.500.750]
      - Ryanodine Receptor Calcium Release Channel [D12.776.210.500.800]
      - Tropomyosin [D12.776.210.500.895]
      - Troponin [D12.776.210.500.910] +

Amino Acids, Peptides, and Proteins [D12]
- Proteins [D12.776]
  - Cytoskeletal Proteins [D12.776.220]
    - Adenomatous Polyposis Coli Protein [D12.776.220.040]
    - Dystrophin [D12.776.220.250]
    - Intermediate Filament Proteins [D12.776.220.475] +
    - Microfilament Proteins [D12.776.220.525] +
    - Microtubule Proteins [D12.776.220.600] +
    - Spectrin [D12.776.220.980]
    - Talin [D12.776.220.985]
    - Vinculin [D12.776.220.990]

Amino Acids, Peptides, and Proteins [D12]
- Proteins [D12.776]
  - Membrane Proteins [D12.776.543]
    - Ankyrins [D12.776.543.080]
    - Arrestins [D12.776.543.090] +
    - Bacterial Outer Membrane Proteins [D12.776.543.100] +
    - Caveolins [D12.776.543.160]
    - Clathrin [D12.776.543.200]
    - Coat Protein Complex I [D12.776.543.212] +
    - Connexins [D12.776.543.225] +
    - Dystrophin [D12.776.543.250]
    - Heterotrimeric GTP-Binding Proteins [D12.776.543.325] +
    - LDL-Receptor Related Protein-Associated Protein [D12.776.543.475]
    - Membrane Glycoproteins [D12.776.543.550] +
    - Membrane Transport Proteins [D12.776.543.585] +
    - Myelin Proteins [D12.776.543.620] +
    - Neurofibromin 2 [D12.776.543.685]
    - Receptors, Cell Surface [D12.776.543.750] +
    - Spectrin [D12.776.543.980]


[UMLS_2002] UMLS® Semantic Navigator ® [2.07] - Netscape 6
File Edit View Search Go Bookmarks Tasks Help
http://etbsun2.nlm.nih.gov:8000/perl/semnav_cgi.pl
Search
Siblings
Chemicals & Drugs
(LA)12 peptide
(methyl)ammonium uptake carrier, Corynebacterium
120-kDa hemocyte-specific membrane protein, flesh fly
15a protein, Aedes aegypti
22.6-kDa antigen, Schistosoma japonicum
36-kDa vesicular integral membrane protein
38L protein
5-lipoxygenase-activatin protein
59 kDa dystrophin-associated protein
A-1 antigen
A-kinase anchor protein 149
A-kinase anchor protein 15
A-kinase anchor protein 200
A-kinase anchor protein KL
A14.5L protein
A15 protein
A4 protein
ABC-me protein
AcfB protein
ACR3 protein
AcrE protein
actA protein
Actinin
Contractile Proteins
Binding Proteins
Microfilament Proteins
proteins by body part
Muscle Proteins
Actin-Binding Proteins
Membrane Proteins
Dystrophin
40-kDa dystrophin
apo-dystrophin 2
dystrophin-related protein 2
Dp260 protein
apo-dystrophin 3
Other Related Concepts
Disorders
Muscular Dystrophies
Muscular Dystrophy, Duchenne
Living Beings
Mice, Inbred mdx
(3 other related concepts)
Co-occurring Concepts
Anatomy
Brain [33]
Cell Membrane [9]
Cytoskeleton [7]
Heart [8]
Hippocampus [7]
Muscle Fibers [32]
Muscle Fibers, Fast-Twitch [5]
Muscle, Skeletal [149]
Muscle, Smooth [6]
Muscle, Smooth, Vascular [6]
Dystrophin
LEGEND
Start again
Apply new parameters
Restrict to vocabulary: Show all
Highlight vocabulary: Nothing
UMLS data: UMLS_2002
Type of hierarchical rel.: All Parent/Child only Broader/Narrower only
Transitive reduction: yes no
Similar Concepts
(none)
Closest MeSH Terms
Main Headings
Dystrophin
Subheadings

SemNav Visualization options



{UMLS_2002} UMLS® Semantic Navigator ¤ [2.07] - Netscape 6

File Edit View Search Go Bookmarks Tasks Help

http://etbsun2.nlm.nih.gov:8000/perl/semnav_cgi.pl

Search

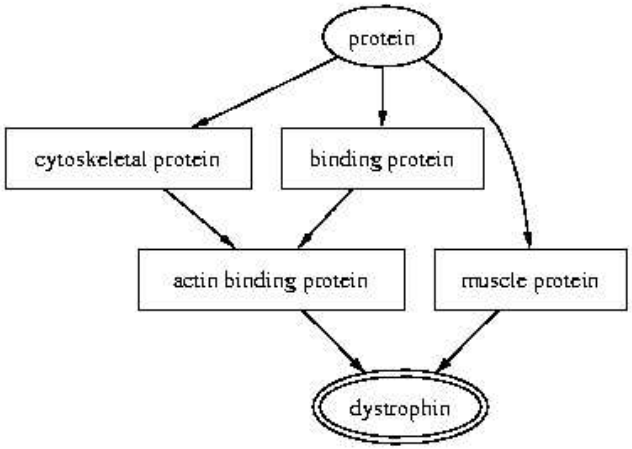
Siblings

Chemicals & Drugs

- alpha actinin ¤
- caldesmon ¤
- calponin ¤
- gelsolin ¤
- muscle relaxing factor ¤
- myoglobin ¤
- myosin ATPase ¤
- synapsin ¤
- tropomyosin ¤
- troponin ¤
- vinculin ¤

(11 siblings)

[children and narrower concepts of parents and broader concepts]



Other Related Concepts

Disorders

- muscular dystrophy ¤

(1 other related concept)

BCI

Dystrophin

LEGEND *

Start again | Apply new parameters

Restrict to vocabulary: CRISP

Highlight vocabulary: Nothing

UMLS data: UMLS_2002

Type of hierarchical rel.: All / Parent/Child only / Broader/Narrower only

Transitive reduction: yes / no

Similar Concepts

(none)

Closest MeSH Terms

Main Headings

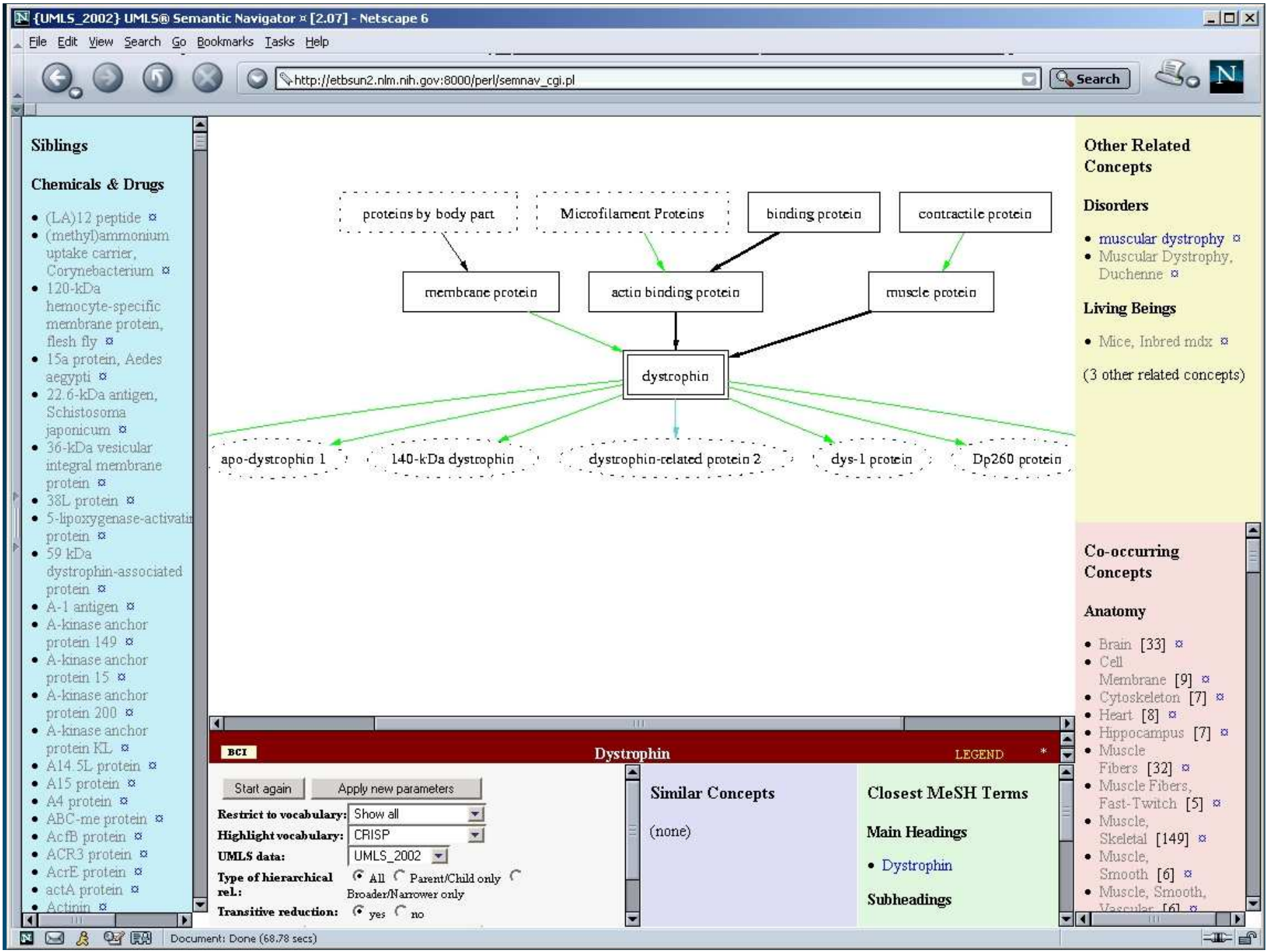
- Dystrophin

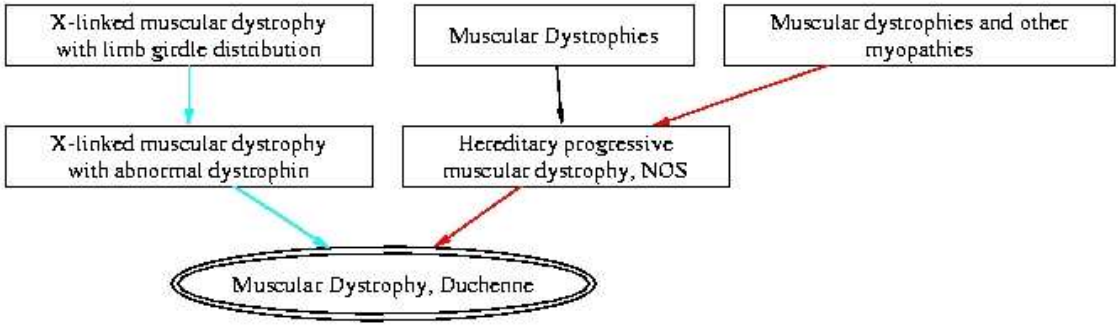
Subheadings

Co-occurring Concepts

(not displayed: Restrict to vocabulary in use)

Document: Done (0.484 secs)


{UMLS_2002} UMLS® Semantic Navigator ¤ [2.07] - Netscape 6
File Edit View Search Go Bookmarks Tasks Help
http://etbsun2.nlm.nih.gov:8000/perl/semnav_cgi.pl
Search
Siblings
Chemicals & Drugs
(LA)12 peptide
(methyl)ammonium uptake carrier, Corynebacterium
120-kDa hemocyte-specific membrane protein, flesh fly
15a protein, Aedes aegypti
22.6-kDa antigen, Schistosoma japonicum
36-kDa vesicular integral membrane protein
38L protein
5-lipoxygenase-activatin protein
59 kDa dystrophin-associated protein
A-1 antigen
A-kinase anchor protein 149
A-kinase anchor protein 15
A-kinase anchor protein 200
A-kinase anchor protein KL
A14.5L protein
A15 protein
A4 protein
ABC-me protein
AcfB protein
ACR3 protein
AcrE protein
actA protein
Actinin
proteins by body part
Microfilament Proteins
binding protein
contractile protein
membrane protein
actin binding protein
muscle protein
dystrophin
apo-dystrophin 1
140-kDa dystrophin
dystrophin-related protein 2
dys-1 protein
Dp260 protein
Other Related Concepts
Disorders
muscular dystrophy
Muscular Dystrophy, Duchenne
Living Beings
Mice, Inbred mdx
(3 other related concepts)
Co-occurring Concepts
Anatomy
Brain [33]
Cell Membrane [9]
Cytoskeleton [7]
Heart [8]
Hippocampus [7]
Muscle Fibers [32]
Muscle Fibers, Fast-Twitch [5]
Muscle, Skeletal [149]
Muscle, Smooth [6]
Muscle, Smooth,
BCI
Dystrophin
LEGEND
Start again
Apply new parameters
Restrict to vocabulary: Show all
Highlight vocabulary: CRISP
UMLS data: UMLS_2002
Type of hierarchical rel.: All Parent/Child only Broader/Narrower only
Transitive reduction: yes no
Similar Concepts
(none)
Closest MeSH Terms
Main Headings
Dystrophin
Subheadings
Document: Done (68.78 secs)



Siblings

Disorders

- Erb's muscular dystrophy
- Gower's muscular dystrophy
- Intermediate X-linked muscular dystrophy
- Manifesting female carrier of X-linked muscular dystrophy
- Muscular Dystrophy, Becker
- Muscular Dystrophy, Facioscapulohumeral
- Muscular Dystrophy, Limb-Girdle
- Muscular Dystrophy, Oculopharyngeal
- Ocular muscular dystrophy

(9 siblings)

[direct children and narrower concepts of direct parents and broader concepts]

Other Related Concepts

Chemicals & Drugs

- Dystrophin

Disorders

- Muscular Dystrophy, Becker
- MUSCULAR DYSTROPHY, PSEUDOHYPERTROPHI PROGRESSIVE, DUCHENNE AND BECKER TYPES
- Progressive muscular dystrophy, legs

(4 other related concepts)

Co-occurring Concepts

Activities & Behaviors

- Knowledge, Attitudes, Practice [1]
- Physician's Practice Patterns [1]
- Psychological adjustment [1]
- Verbal Behavior [1]

Anatomy

- Achilles Tendon [1]
- Ankle [1]

Muscular Dystrophy, Duchenne

LEGEND

Start again | Apply new parameters

Restrict to vocabulary: Show all

Highlight vocabulary: Nothing

UMLS data: UMLS_2002

Type of hierarchical rel.: All / Parent/Child only / Broader/Narrower only

Transitive reduction: yes / no

Similar Concepts

- Hereditary progressive muscular dystrophy, NOS

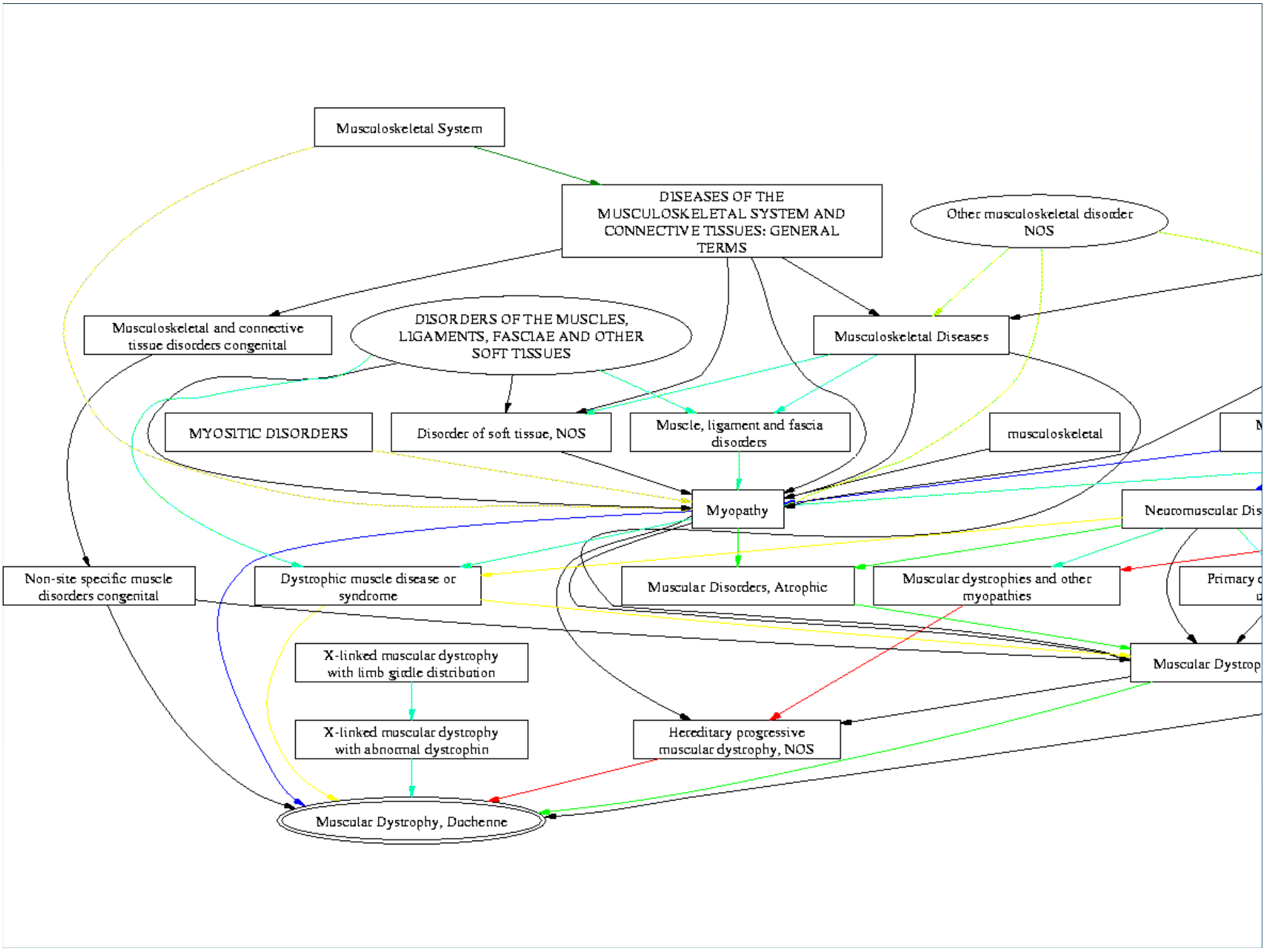
(1 similar concept)

Closest MeSH Terms

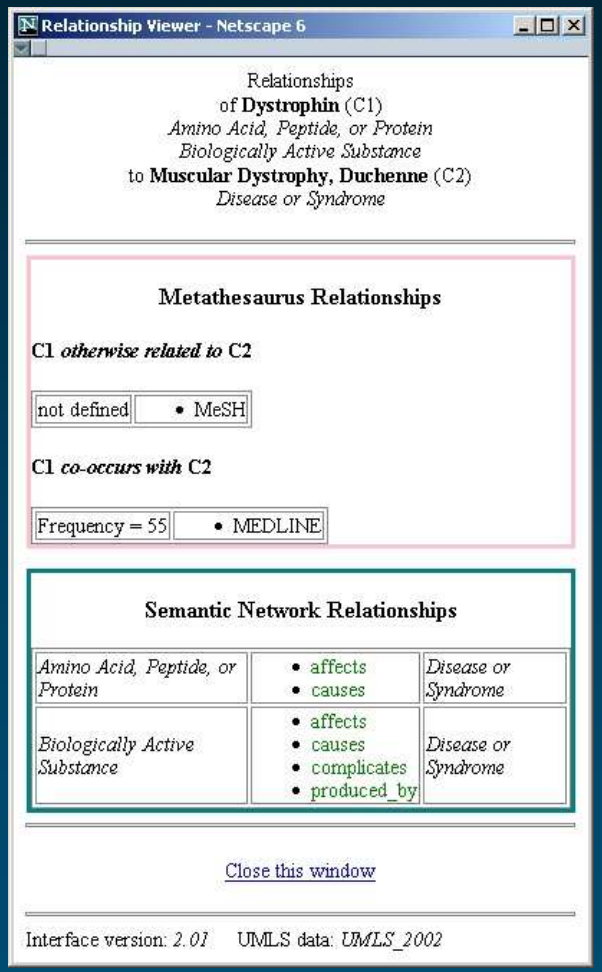
Main Headings

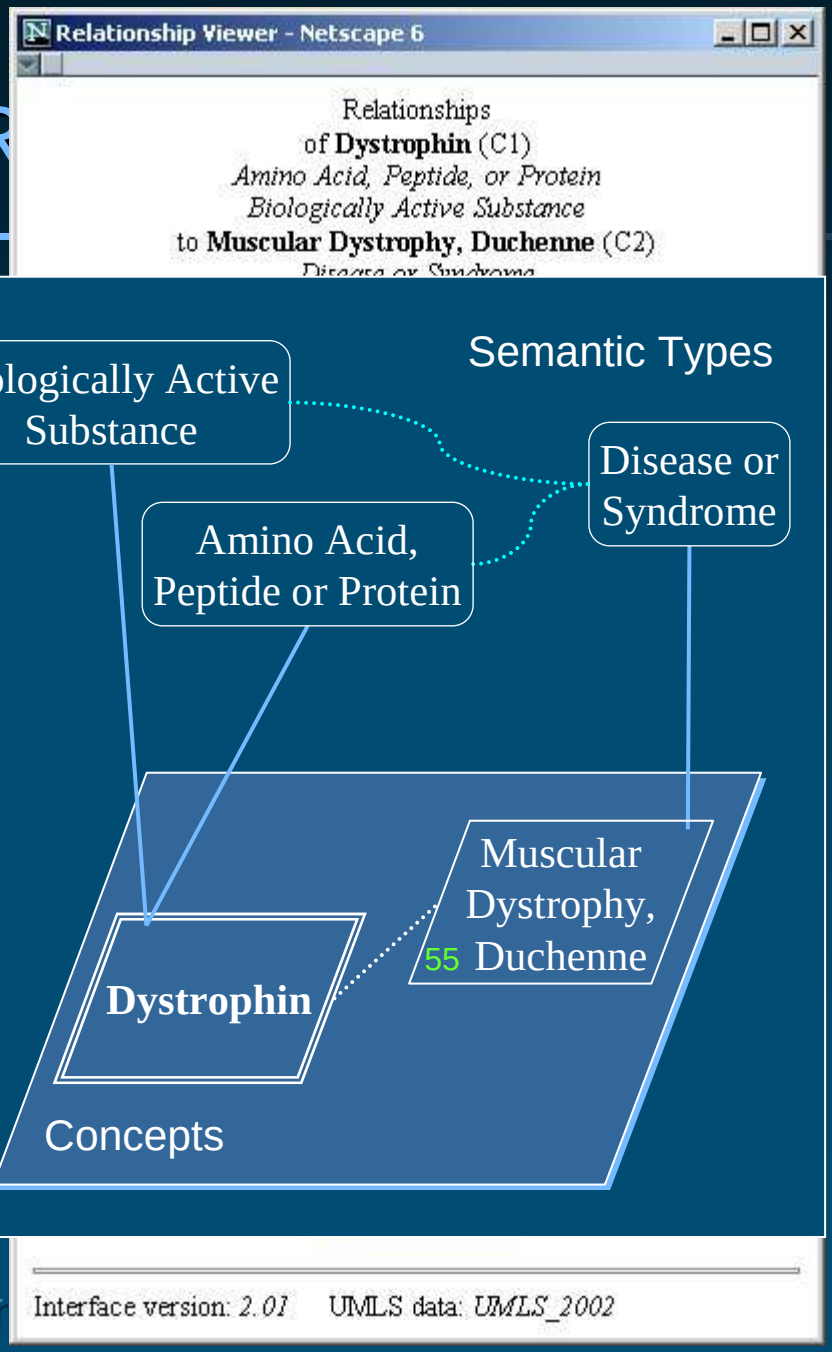
- Muscular Dystrophy, Duchenne

Subheadings


Musculoskeletal System
DISEASES OF THE MUSCULOSKELETAL SYSTEM AND CONNECTIVE TISSUES: GENERAL TERMS
Other musculoskeletal disorder NOS
Musculoskeletal and connective tissue disorders congenital
DISORDERS OF THE MUSCLES, LIGAMENTS, FASCIAE AND OTHER SOFT TISSUES
Musculoskeletal Diseases
MYOSITIC DISORDERS
Disorder of soft tissue, NOS
Muscle, ligament and fascia disorders
musculoskeletal
Myopathy
Neuromuscular Dis
Non-site specific muscle disorders congenital
Dystrophic muscle disease or syndrome
Muscular Disorders, Atrophic
Muscular dystrophies and other myopathies
Primary
X-linked muscular dystrophy with limb girdle distribution
Muscular Dystrop
X-linked muscular dystrophy with abnormal dystrophin
Hereditary progressive muscular dystrophy, NOS
Muscular Dystrophy, Duchenne

SemNav R


Relationship Viewer - Netscape 6
Relationships
of Dystrophin (C1)
Amino Acid, Peptide, or Protein
Biologically Active Substance
to Muscular Dystrophy, Duchenne (C2)
Disease or Syndrome
Metathesaurus Relationships
C1 otherwise related to C2
not defined
• MeSH
C1 co-occurs with C2
Frequency = 55
• MEDLINE
Semantic Network Relationships
Amino Acid, Peptide, or Protein
• affects
• causes
Disease or Syndrome
Biologically Active Substance
• affects
• causes
• complicates
• produced_by
Disease or Syndrome
Close this window
Interface version: 2.01 UMLS data: UMLS_2002

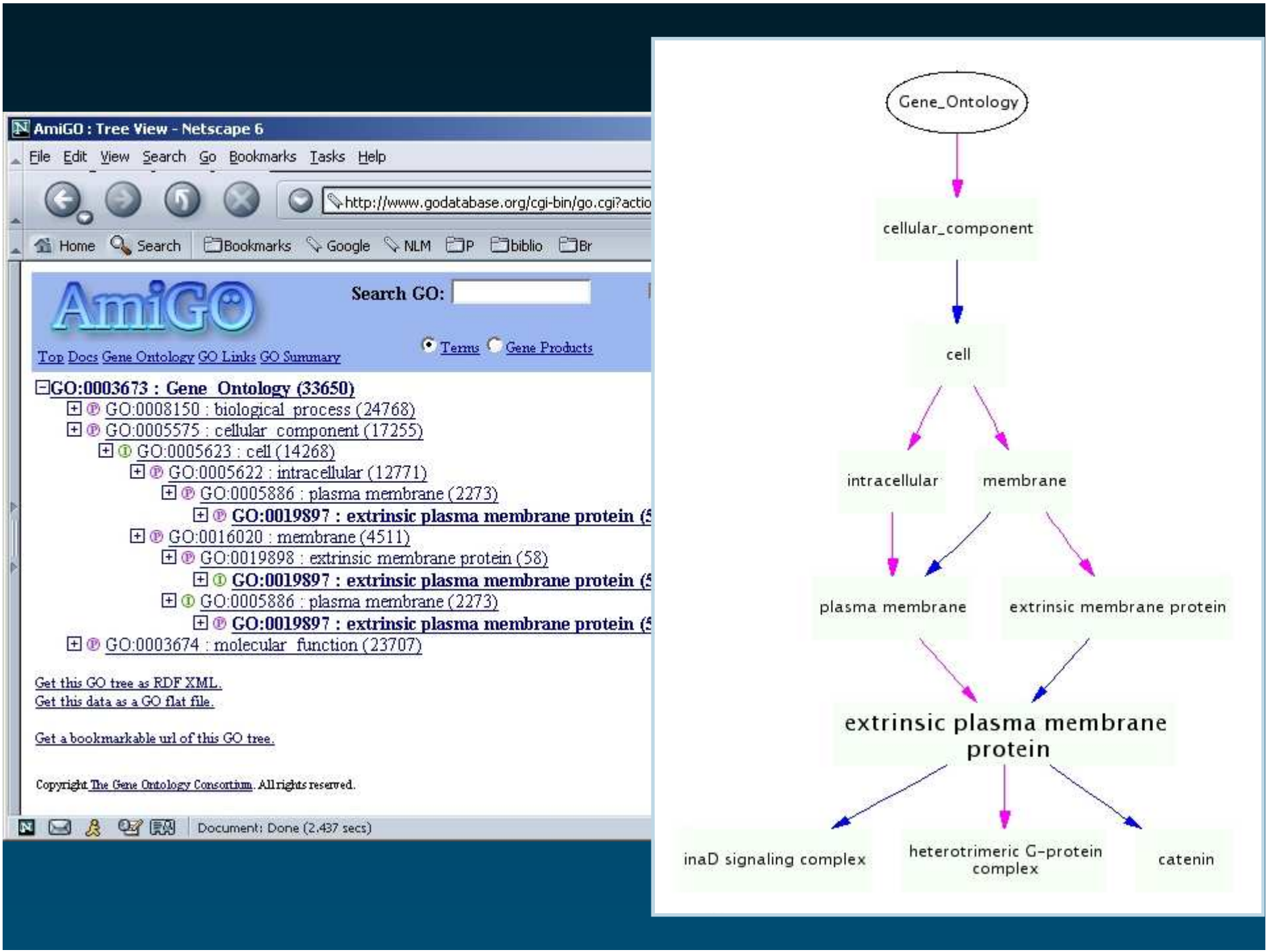

Relationship Viewer - Netscape 6
Relationships
of Dystrophin (C1)
Amino Acid, Peptide, or Protein
Biologically Active Substance
to Muscular Dystrophy, Duchenne (C2)
Semantic Types
Biologically Active Substance
Disease or Syndrome
Amino Acid, Peptide or Protein
Muscular Dystrophy, Duchenne
55
Dystrophin
Concepts
Interface version: 2.01 UMLS data: UMLS_2002

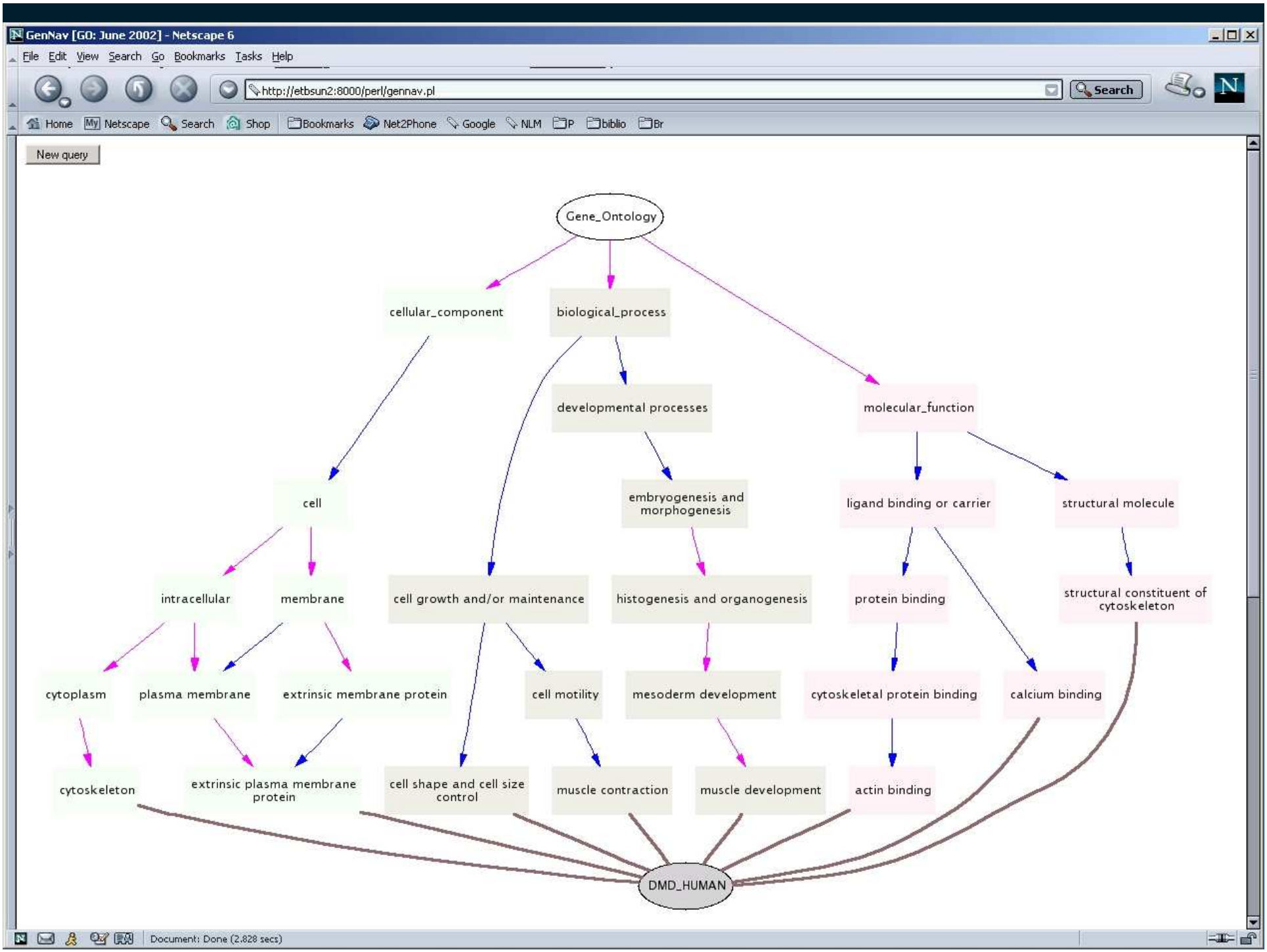
Gene Ontology browser

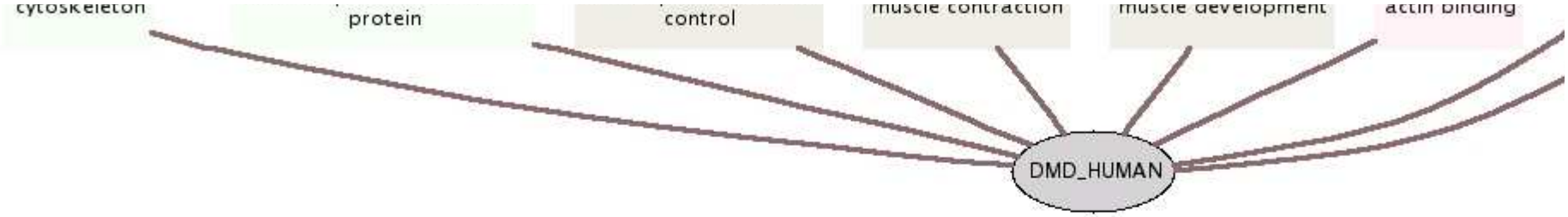
http://mor.nlm.nih.gov/perl/gennav.pl

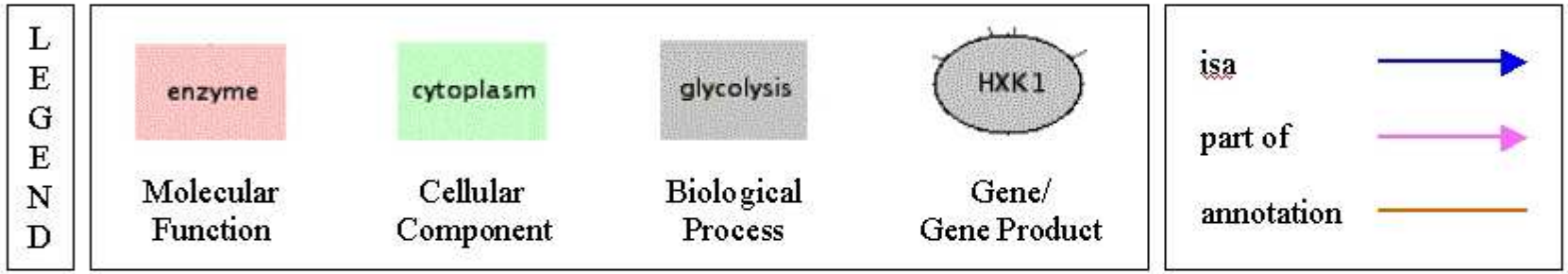
Gene Ontology™

- Developed by the GO Consortium
- Several components (GO database)
  - Ontology (~17,000 concepts)
    - Molecular functions
    - Cellular components
    - Biological processes
  - Gene products (~1.6M)
  - Associations between Gene products and GO concepts (~6.8M)


AmiGO : Tree View - Netscape 6
File Edit View Search Go Bookmarks Tasks Help
http://www.godatabase.org/cgi-bin/go.cgi?actio
Home Search Bookmarks Google NLM P biblio Br
AmiGO
Search GO:
Top Docs Gene Ontology GO Links GO Summary
Terms Gene Products
GO:0003673 : Gene_Ontology (33650)
GO:0008150 : biological_process (24768)
GO:0005575 : cellular_component (17255)
GO:0005623 : cell (14268)
GO:0005622 : intracellular (12771)
GO:0005886 : plasma membrane (2273)
GO:0019897 : extrinsic plasma membrane protein (5
GO:0016020 : membrane (4511)
GO:0019898 : extrinsic membrane protein (58)
GO:0019897 : extrinsic plasma membrane protein (5
GO:0005886 : plasma membrane (2273)
GO:0019897 : extrinsic plasma membrane protein (5
GO:0003674 : molecular_function (23707)
Get this GO tree as RDF XML.
Get this data as a GO flat file.
Get a bookmarkable url of this GO tree.
Copyright The Gene Ontology Consortium. All rights reserved.
Document: Done (2.437 secs)
Gene_Ontology
cellular_component
cell
intracellular
membrane
plasma membrane
extrinsic membrane protein
extrinsic plasma membrane protein
inaD signaling complex
heterotrimeric G-protein complex
catenin


GenNav [GO: June 2002] - Netscape 6
File Edit View Search Go Bookmarks Tasks Help
http://etbsun2:8000/perl/gennav.pl
Search
Home Netscape Search Shop Bookmarks Net2Phone Google NLM P biblio Br
New query
Gene_Ontology
cellular_component
biological_process
developmental processes
molecular_function
cell
embryogenesis and morphogenesis
ligand binding or carrier
structural molecule
intracellular
membrane
cell growth and/or maintenance
histogenesis and organogenesis
protein binding
structural constituent of cytoskeleton
cytoplasm
plasma membrane
extrinsic membrane protein
cell motility
mesoderm development
cytoskeletal protein binding
calcium binding
cytoskeleton
extrinsic plasma membrane protein
cell shape and cell size control
muscle contraction
muscle development
actin binding
DMD_HUMAN
Document: Done (2.828 secs)



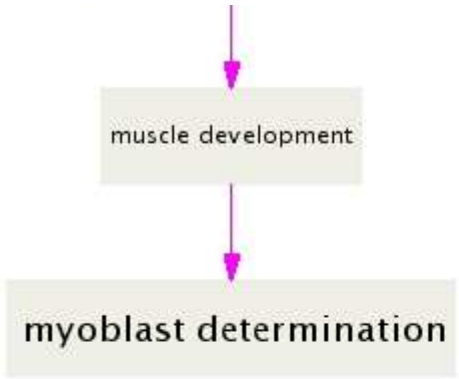


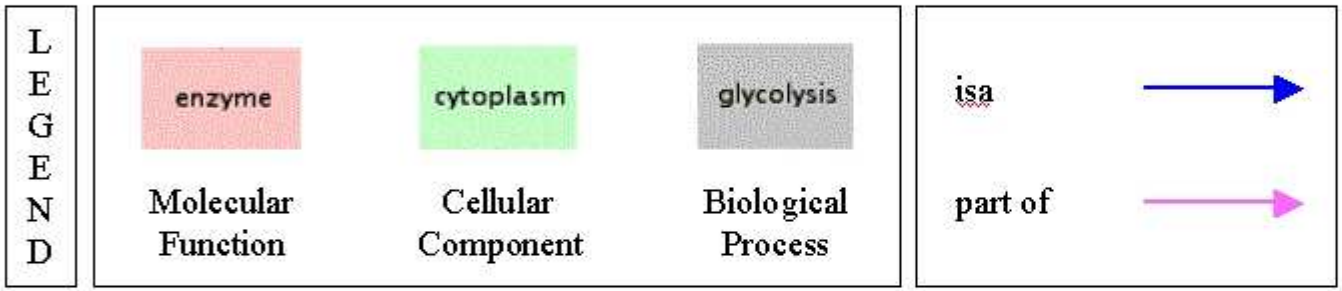
Symbol	DMD_HUMAN
Species	0
Full name	Dystrophin
Synonyms	• IPI00006091
Cross-references	DB: sp; TYPE: acc; ID: P11532; OBJ_TYPE: protein; TAXON: 9606
Molecular functions	• actin binding [IEA] • structural constituent of cytoskeleton [TAS] • calcium binding [IEA]
Biological processes	• muscle contraction [NR] • cell shape and cell size control [TAS] • muscle development [NR]
Cellular components	• extrinsic plasma membrane protein [TAS] • cytoskeleton [TAS]

New query

Interface version: *0.3*

Document: Done (0.593 secs)



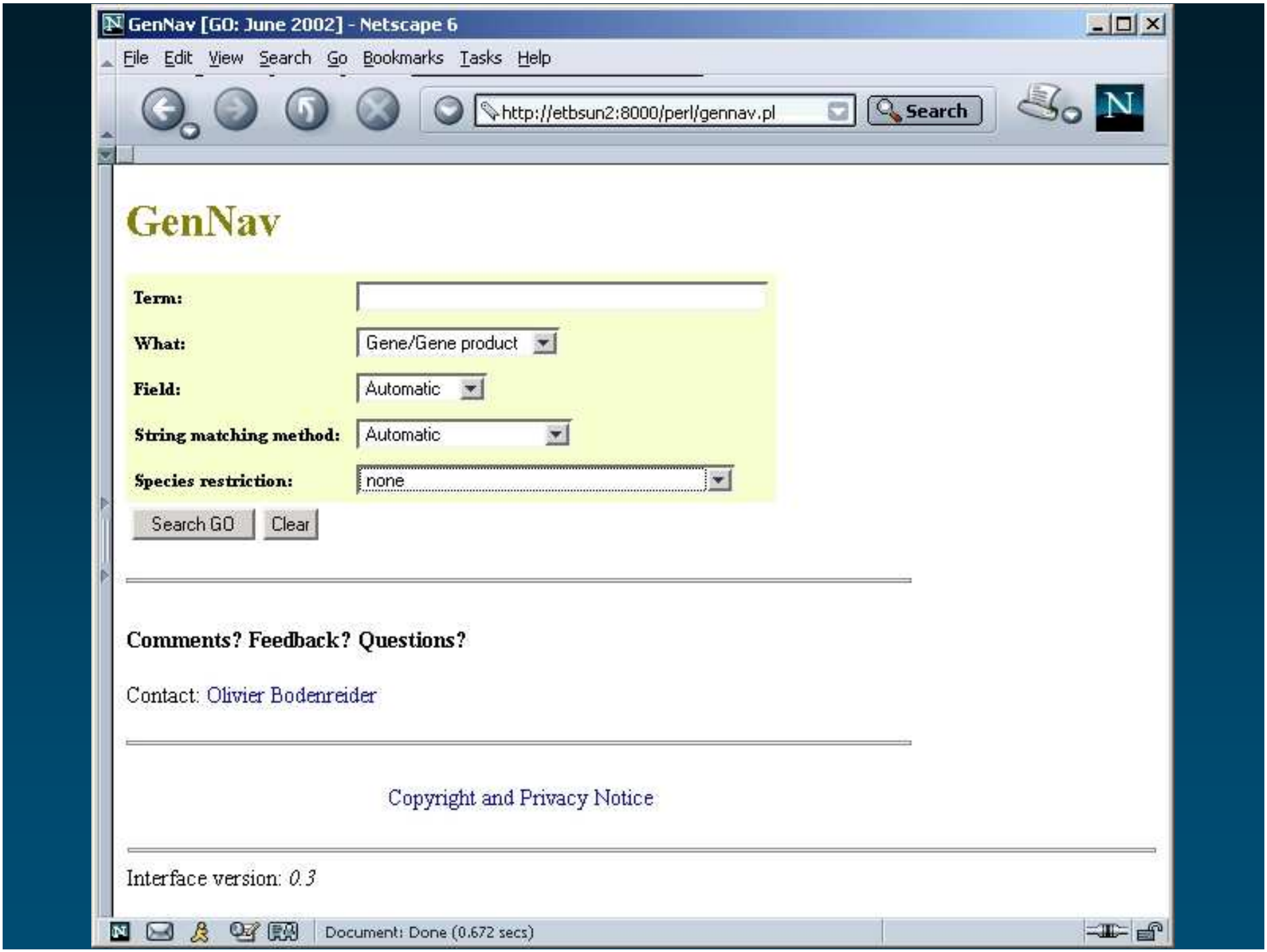


Name	myoblast determination
Type	process
Accession number	7518
Synonyms	none
Definitions	none
Parents (isa)	none
Parents (part of)	• muscle development
Children (isa)	none
Children (part of)	none
Cross-references	ID KEY DB

Gene products (restricted to species: 9606)

Full name	Symbol	Taxon	Evidence
Interferon-related developmental regulator 1	IFR1_HUMAN	9606	TAS
MEF-2 PROTEIN	Q9Y655	9606	TAS

New query


GenNav [GO: June 2002] - Netscape 6
File Edit View Search Go Bookmarks Tasks Help
http://etbsun2:8000/perl/gennav.pl
Search
GenNav
Term:
What:
Gene/Gene product
Field:
Automatic
String matching method:
Automatic
Species restriction:
none
Search GO
Clear
Comments? Feedback? Questions?
Contact: Olivier Bodenreider
Copyright and Privacy Notice
Interface version: 0.3
Document: Done (0.672 secs)

Technical details

Technical details

- Simple web/cgi technology (apache, Perl)
- dot (GraphViz)
  - PNG file (-Tpng)
  - Client-side map (-Tcmap)
- Precompute the transitive closure on hierarchical relations to perform the transitive closure fast
- Remove cycles (UMLS)

Discussion

Issues and Challenges

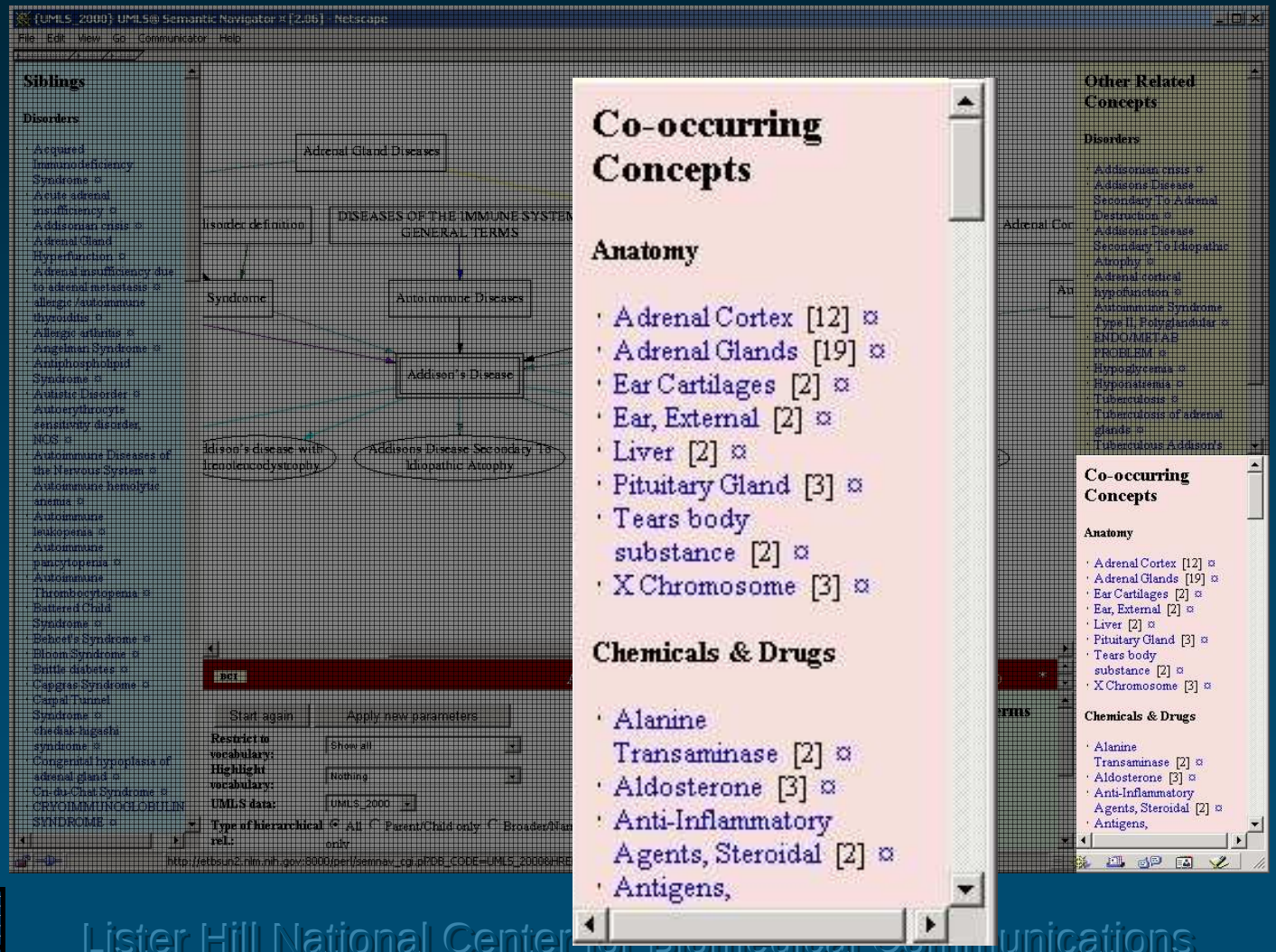
Issues

- Size
  - Large number of concepts (>1 million)
- Complexity
  - Polyhierarchical structures
  - Multiple information sources
  - Multiple properties
- Lack of formality
  - Redundant relations
  - Hierarchies vs. hierarchical relations

Challenges

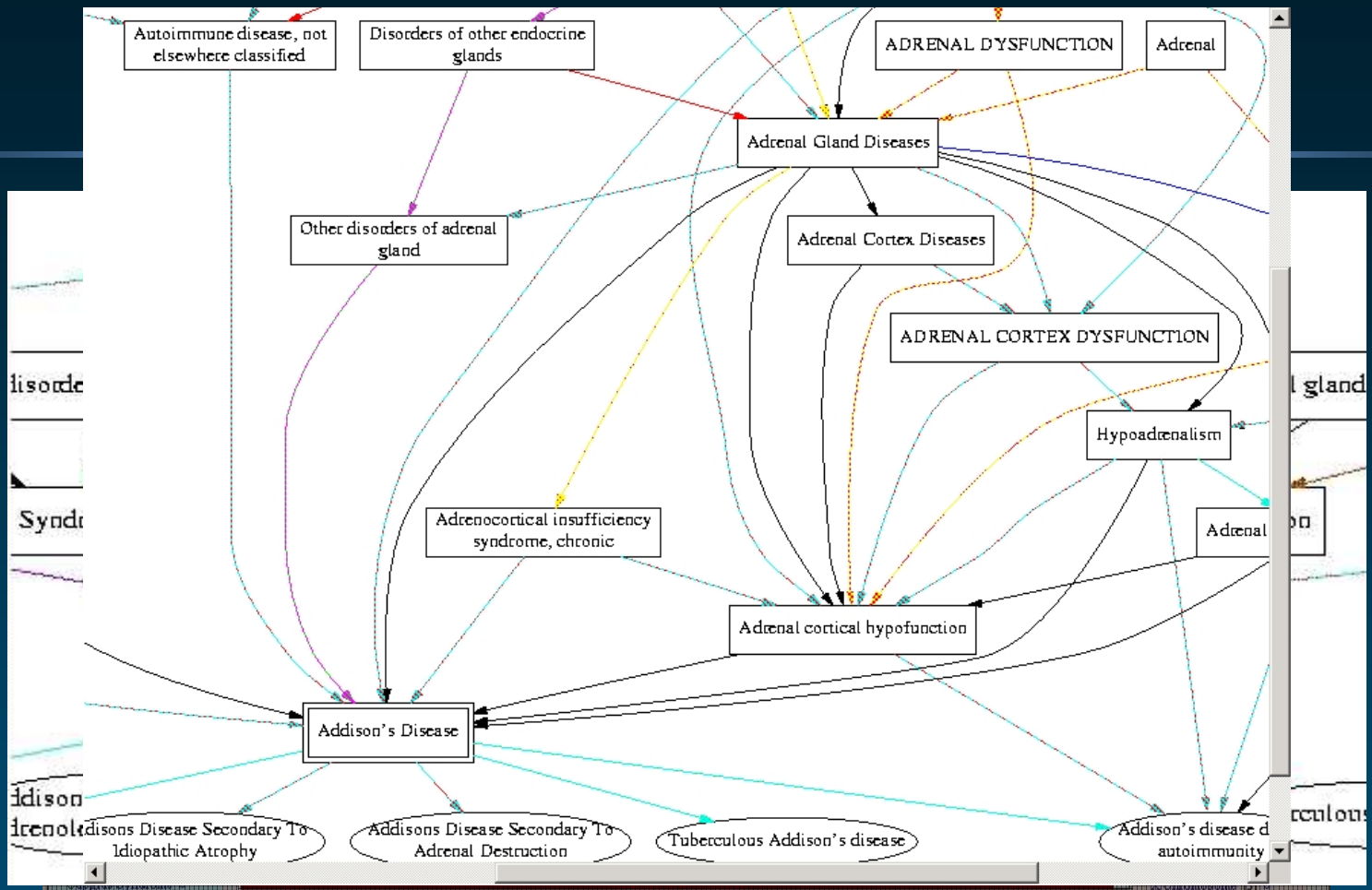
- Restrict information space
  - To selected information sources (SemNav)
  - To selected organisms (GenNav)
- Reduce complexity (SemNav)
  - Group concepts by semantic groups
  - Transitive reduction on hierarchical relations
  - Select co-occurring concepts
- Reduce the cognitive burden on the user
  - Use graph-based rather than tree-based representations

SemNav Semantic groups



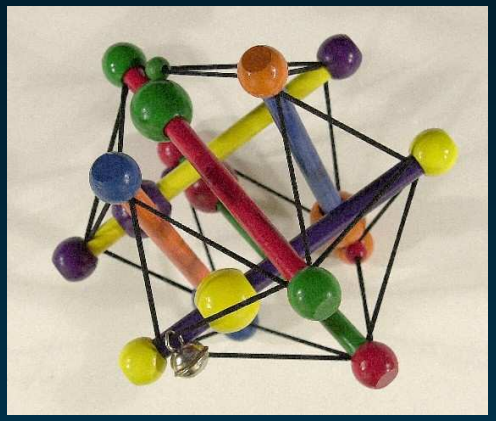
Challenges

- Restrict information space
  - To selected information sources (SemNav)
  - To selected organisms (GenNav)
- Reduce complexity (SemNav)
  - Group concepts by semantic groups
  - Transitive reduction on hierarchical relations
  - Select co-occurring concepts
- Reduce the cognitive burden on the user
  - Use graph-based rather than tree-based representations


Autoimmune disease, not elsewhere classified
Disorders of other endocrine glands
ADRENAL DYSFUNCTION
Adrenal
Adrenal Gland Diseases
Other disorders of adrenal gland
Adrenal Cortex Diseases
ADRENAL CORTEX DYSFUNCTION
Hypoadrenalism
Adrenocortical insufficiency syndrome, chronic
Adrenal
Adrenal cortical hypofunction
Addison's Disease
Addisons Disease Secondary To Idiopathic Atrophy
Addisons Disease Secondary To Adrenal Destruction
Tuberculous Addison's disease
Addison's disease d autoimmunity

Challenges

- Restrict information space
  - To selected information sources (SemNav)
  - To selected organisms (GenNav)
- Reduce complexity (SemNav)
  - Group concepts by semantic groups
  - Transitive reduction on hierarchical relations
  - Select co-occurring concepts
- Reduce the cognitive burden on the user
  - Use graph-based rather than tree-based representations



Medical Ontology Research

Contact: olivier@nlm.nih.gov

Web: mor.nlm.nih.gov

Olivier Bodenreider

Lister Hill National Center
for Biomedical Communications
Bethesda, Maryland - USA